REPORT DOCUMENTATION PAGE

1

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
12-09-2014	Final Report	1-Sep-2010 - 31-Aug-2014

4. TITLE AND SUBTITLE
Final Report: Networking in the Presence of Adversaries

5a. CONTRACT NUMBER
W911NF-10-1-0419

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Lang Tong

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Cornell University
373 Pine Tree Road

Ithaca, NY 14850 -2820

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
58094-NS.17

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
This research addresses tradeoffs between reliability and efficiency of communication networks when the network is subject to attacks by adversaries. The overall objective of this research is twofold. First, this research aims to develop a mathematical theory that characterizes fundamental limits of networking in the presence of inside (Byzantine) adversaries. To this end, the maximum rate of reliable communication is quantified as a function of the number of covert adversarial nodes. Second, this research develops practical coding schemes that guarantee reliable communication (in the information theoretic sense) in the presence of arbitrary attacks by intelligent adversaries

15. SUBJECT TERMS
communication networks, security, information theory, network coding

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lang Tong
UU	UU	UU	UU		19b. TELEPHONE NUMBER 607-255-3900

Report Title

Final Report: Networking in the Presence of Adversaries

ABSTRACT

This research addresses tradeoffs between reliability and efficiency of communication networks when the network is subject to attacks by adversaries. The overall objective of this research is twofold. First, this research aims to develop a mathematical theory that characterizes fundamental limits of networking in the presence of inside (Byzantine) adversaries. To this end, the maximum rate of reliable communication is quantified as a function of the number of covert adversarial nodes. Second, this research develops practical coding schemes that guarantee reliable communication (in the information theoretic sense) in the presence of arbitrary attacks by intelligent adversaries. Both source and channel coding aspects are investigated. Extensions to cyber-physical systems are also considered.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received		Paper
09/09/2014	7.00	Oliver Kosut, Lang Tong, David N. C. Tse. Polytope Codes Against Adversaries in Networks, IEEE Transactions on Information Theory, (06 2014): 0. doi: 10.1109/TIT.2014.2314642
09/09/2014	9.00	Brandon M. Jones, Mark Campbell, Lang Tong. Maximum Likelihood Fusion of Stochastic Maps, IEEE Transactions on Signal Processing, (04 2014): 0. doi: 10.1109/TSP.2014.2304435
10/23/2012	3.00	P. Venkitasubramaniam, Lang Tong. A Game-Theoretic Approach to Anonymous Networking, IEEE/ACM Transactions on Networking, (06 2012): 892. doi: 10.1109/TNET.2011.2176511
TOTAL:	**3**	

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received		Paper
TOTAL:		

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper	
09/09/2014	13.00	Lang Tong, Jinsub Kim, Robert J. Thomas. Data framing attack on state estimation with unknown network parameters, 2013 Asilomar Conference on Signals, Systems and Computers. 03-NOV-13, Pacific Grove, CA, USA. : ,
09/09/2014	14.00	Xiaoqing Fan, Aaron B. Wagner, Ebad Ahmed. Polytope codes for large-alphabet channels, 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton). 02-OCT-13, Monticello, IL. : ,
10/23/2011	1.00	Mark Campbell, Lang Tong, Brandon Jones. Maximum Likelihood Combining of Stochastic Maps, 2011 Allerton Conference on Communication, Control and Computing,. 29-SEP-11, . : ,
10/23/2011	2.00	Oliver Kosut, Lang Tong, David Tse. Polytope Codes Against Adversaries in Networks, 2010 Intl Symp. Inform. Theory. , . : ,
10/23/2012	4.00	Ebad Ahmed, Aaron Wagner. Lossy Source Coding with Byzantine, 2011 IEEE Information Theory Workshop (ITW). 30-JUN-11, . : ,
TOTAL:	**5**	

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received		Paper
09/09/2014	8.00	Jinsub Kim, Lang Tong, Robert J. Thomas. Data Framing Attack on State Estimation, IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS (04 2014)
09/09/2014	15.00	Jinsub Kim, Lang Tong, Robert J. Thomas. Subspace methods for data attack on state estimation: a data driven approach, IEEE TRANSACTIONS ON Signal Processing (05 2014)
10/23/2012	6.00	Oliver Kosut, Lang Tong, David Tse. Polytope codes against adversaries in networks, IEEE Transactions in Information Theory (12 2011)
TOTAL:		**3**

Number of Manuscripts:

Books

Received	Book
TOTAL:	

Received	Book Chapter
TOTAL:	

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Oliver Kosut	0.20	
Ebad Ahmed	0.20	
Jinsub Kim	0.40	
Kathy Fan	0.10	
Brandon Jones	0.10	
FTE Equivalent:	**1.00**	
Total Number:	**5**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Lang Tong	0.15	
FTE Equivalent:	**0.15**	
Total Number:	**1**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:..... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):..... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:..... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: 0.00

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME	
Oliver Kosut	
Jinsub Kim	
Ebad Ahmed	
Brandon Jones	
Total Number:	**4**

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Since the seminal work of Shannon in 1948, reliability and efficiency tradeoff have been at the heart of modern communication theory. For random errors introduced in the communication channels, Shannon's information theory shows where the tradeoff between reliability and efficiency lies, and information theory has inspired practical coding schemes for detecting and correcting such random errors.

What happens when errors introduced are not random effects of nature but results of malicious acts of an adversary? As modern communications rely increasingly on networks, what if some of the nodes are controlled by adversaries? These questions, unfortunately, cannot be answered easily using classical information and coding theories because actions of an adversary cannot be captured by some probability distributions. More importantly, adversaries can be cognitive and opportunistic as they learn the communication and networking environments and adapt their strategies of attacks. Despite the rapid expansion of literature on communication and network security, we know very little at the fundamental level how to characterize the reliability-efficiency tradeoff when there are adversaries in the network, and we lack provably effective techniques to mitigate malicious and covert actions inside the network.

If communication network is vulnerable to attacks, cyber physical systems (CPS) that rely on communication networks for sensing, actuation, and control will also vulnerable. Because CPS often involves coordinated decisions for real-time operations, adversarial attacks on CPS carry significantly different objectives from gaining access to information; such attacks may aim at disrupting critical mission.

This project addresses security issues of communication networks and, more broadly, cyber physical systems are subject to internal or external attacks. The overall objective of this research is three. First, we develop a mathematical theory that characterizes limits of networking and fundamental tradeoffs among networking performance (capacity, throughput, delay, etc.), measures of security (e.g., probability of detection), and the power of adversaries (the number of channels that adversary can monitor, the number of adversarial nodes). Second, we develop schemes for practical applications, which include new coding techniques capable of countering arbitrary adversarial actions. Finally, we extend theory developed here to broader classes of networks. In particular, we focus on cyber-physical systems where the aims of attack go beyond creating decoding error.

The key approach considered in this research is to develop the class of structured nonlinear codes. To this end, we have investigated the class of nonlinear codes that are built upon the classical structures of MDS (linear) codes with additional anomaly detection capabilities. This latter feature is crucial to discover adversary actions not at the destination but at an earlier stage of the information flow.

The project has also addressed important architectural questions. In particular, we have examined the applicability of various "separation principles" in network design in the presence of adversaries and identify scenarios when designs based on separation principles are optimal and to what degree suboptimal.

Beside communication networks, we have considered cyber physical systems under the so-called man-in-the middle attack, focusing on attacks on state estimation, which is the key component of any CPS. The mathematical abstract considered in our work arises from electric networks that are monitored by sensors. However, the attack and countermeasure schemes developed apply to more general settings.

Significance

The tactical networks for the military must operate in hostile environments where nodes of the network are vulnerable to adversary attacks. As operations increasingly rely on networked distributed systems, the risk of attacks also increases. The advent of P2P operations, cloud computing, network coding, and sensor networks for military tactical networking raises cogent needs of developing coding and networking mechanisms that provide reliable operations in the presence of unreliable and possibly adversarial participants.

The results obtained provide important insights into performance cost in the presence of adversaries. They illustrate how mathematical structures of network coding and lossy compression schemes. Of particular significance is the implication on the source-channel separation principle, which has been shown to be optimal in a wide range of scenarios in conventional networks, might not in general be optimal when adversaries are present. This fact is particularly intriguing and indicates that the fundamental limits of adversarial networks might be appreciably different from those of conventional networks.

Summary of the most important results

1.Polytope Codes Against Adversaries in Networks
Network coding allows routers in a network to execute possibly complex codes in addition to routing; it has been shown that allowing them to do so can increase communication rate. However, taking advantage of coding at internal nodes means that

sources and destinations must rely on other nodes—nodes they may not have complete control over—to reliably perform certain functions. If these internal nodes do not behave correctly, or, worse, maliciously attempt to subvert the goals of the users—constituting a so-called Byzantine attack—standard network coding techniques fail.

Our primary contribution is a class of network codes to defeat adversaries called Polytope Codes. These were originally introduced in under the less descriptive term “bounded-linear codes”. Polytope Codes are nonlinear codes, and they improve over linear codes by allowing error detection inside the network. This allows adversaries to be more easily identified, whereby the messages they send can be ignored. We also prove a cut-set upper bound on achievable rates in networks with node-based adversaries. This cut-set bound is a form of the Singleton bound, originally proved for classical error-correcting codes. We show that for a class of planar networks, Polytope Codes can achieve the rate given by this cut-set bound, which means that they achieve the capacity for these networks. We also show that the cut-set bound is not always achievable, by giving an example network with a strictly smaller capacity.

2.Lossy Source Coding with Byzantine Adversaries
While the rapid growth of modern-day communication networks makes them increasingly useful, it also makes them increasingly difficult to protect against attacks. This is especially true of those networks, such as peer-to-peer systems,in which the nodes are controlled by different entities. In the case of peer-to peer networks, malicious users could sabotage the file-sharing process by intentionally transmitting a corrupted version of the file. Similar problems can potentially arise in ad-hoc networks and distributed storage systems.

There has been considerable work on how to protect transmitted information against malicious users within the context of channel- and network-coding, and a number of significant results are available. Yeung and Cai show that if z unit-capacity edges in an acyclic multicast network are subject to random or adversarial errors, then the network capacity is C-2z, where C is the network capacity when all edges are error-free. Thus if an adversary controls z edges, it effectively removes 2z edges from the original adversary-free network. This is reminiscent of the Singleton bound, and we refer to it as the “factor-of-2” rule. The factor-of-2 rule was also shown to hold for lossless source coding: it is well known that if a source X is to be losslessly communicated via n packets, then the sum rate of those packets must be at least the entry of X, H(X). Kosut and Tong have shown that if t of the n packets can be altered in arbitrary ways by adversaries, then every n-2t packets must have sum rate at least H(X). Thus t traitors effectively remove 2t packets from the original adversary-free problem, i.e., the factor-of-2 rule obtains. In the context of peer-to-peer systems, often the ultimate goal is to communicate a file approximately rather than reliably. Codes and fundamental limits for this problem are less well understood. One natural approach to this problem is to perform separate compression and adversarial error-protection. That is, one combines rate-distortion-optimal lossy compression with network codes that are optimal for the adversarial model at hand.

We show that this approach is optimal in some cases but suboptimal in general, even for networks with one sender, one receiver, and no intermediate nodes. Specifically, we consider the problem in which a source is compressed to form n packets, any t of which can be altered in an arbitrary way. The decoder receives the n packets and, without knowing which packets were altered, must estimate the source to meet a given distortion constraint. We show that separate compression and adversarial error correction achieve rate-distortion performance governed by the factor-of-2 rule, and that this is optimal for binary sources with the Hamming distortion measure and Gaussian sources with the mean square error distortion measure. These two optimality results hinge on a combinatorial result of Kleitman on the maximum size of subsets of Hamming space with a given diameter, and the Brunn-Minkowski inequality, respectively. We then show by means of a counterexample, involving a binary source with erasure distortion, that separation is not optimal in general. We consider a 3-encoder problem with one traitor such that one encoder has rate R < 1, while the other two have rate 1 and can therefore transmit the source sequence exactly. We determine the optimal distortion for this problem as a function of R and show that separation cannot achieve it. We note that while source-channel separation has long been known to fail in many scenarios, the reason that it fails here seems to be fundamentally different from the standard examples.

3.Man-in-the-middle Attacks on Cyber Physical Systems
We consider the problem of man-in-the-middle (MiM) attacks on the state estimation of a cyber physical system (CPS) modeled by a topological graph with linear algebraic constraints. As a practical example, such a model arises from an electric power system in which the power flow is governed by the Kirchhoff law. When an adversary launches an MiM data attack, part of the sensor data are intercepted and substituted with malicious data in such a way that the state estimator yields possibly drastically wrong estimates.

We show that if an adversary has the ability to adjust the measurements from enough meters, then no algorithm at the control center will ever be able to detect that an adjustment has been made. This can be viewed as a fundamental limit on the ability of the classical formulation of state estimation to handle cooperative attacks. We also show that there is a close relationship between the attacks described in and system observability. For this reason, we refer to the attacks of as unobservable attacks. This relationship allows us to extend earlier topological results to give an efficient algorithm to calculate attacks of this nature require a small number of adversarial meters. Our algorithm is based on the special structure of the power system, and makes use of techniques to efficiently minimize submodular functions. Our algorithms allow an operator of a power system to find the places in which it is most vulnerable to these attacks.

We consider the problem of MiM attacks on network topology and state estimates. We characterize conditions under which undetectable attacks are possible, given a set of vulnerable meters that may be controlled by an adversary. To this end, we consider two attack regimes based on the information set available to the attacker. The more information the attacker has, the stronger its ability to launch a sophisticated attack that is hard to detect.

The global information regime is where the attacker can observe all meter and network data before altering the adversary-controlled part of them. Although it is unlikely in practice that an adversary is able to operate in such a regime, in analyzing the impact of attacks, it is typical to consider the worst case by granting the adversary additional power. We present a necessary and sufficient algebraic condition under which, given a set of adversary controlled meters, there exists an undetectable attack that misleads the control center with an incorrect “target” topology. This algebraic condition provides not only numerical ways to check if the grid is vulnerable to undetectable attacks but also insights into which meters to protect to defend against topology attacks. We also provide specific constructions of attacks and show certain optimality of the proposed attacks.

A more practically significant situation is the local information regime where the attacker has only local information from those meters it has gained control. We present that, under certain conditions, undetectable attacks exist and can be implemented easily based on simple heuristics. Second, we study conditions under which any topology attack can be made detectable. Such a condition, even if it may not be the tightest, provides insights into defense mechanisms against topology attacks. We show that if a set of meters satisfying a certain branch covering property are protected, then topology attacks can always be detected.

Finally, we consider the data framing attacks that cause the misidentification of good data as bad. Specifically, we formulate the design of optimal data framing attack as a quadratically constrained quadratic program (QCQP). To analyze the efficacy of the data framing attack, we present a sufficient condition under which the framing attack can achieve an arbitrary perturbation of the state estimate by controlling only half of the critical set of meters. The optimal design of framing attack is based on a linearized system. In practice, a nonlinear state estimator is often used. We demonstrate that, under the nonlinear measurement model, the framing attacks designed based on linearized system model successfully perturb the state estimate, and the adversary can control the degree of perturbation as desired.

Technology Transfer